

# Effective Hamiltonians with Relativistic Corrections

*I) The Foldy–Wouthuysen transformation versus the direct Pauli reduction*

H. W. Fearing, G. I. Poulis and S. Scherer  
TRIUMF, Vancouver, British Columbia, Canada V6T 2A3

## Abstract

Two different methods of obtaining “effective  $2 \times 2$  Hamiltonians” which include relativistic corrections to nonrelativistic calculations are discussed. The standard Foldy–Wouthuysen transformation generates Hamiltonians which order by order in  $1/M$  decouple the upper from the lower components. The upper left–hand block then defines an effective  $2 \times 2$  Foldy–Wouthuysen Hamiltonian. In the second method the matrix element of the interaction Hamiltonian of the Dirac representation is evaluated between free positive–energy states and reduced to two–component form. The resulting expression (possibly expanded in  $1/M$ ) then defines what we call the “direct Pauli reduction” effective  $2 \times 2$  Hamiltonian. We wish to investigate under which circumstances the two approaches yield the same result. Using a generic interaction with harmonic time dependence we show that differences in the corresponding effective S–matrices do arise beyond first–order perturbation theory. We attribute them to the fact that the use of the direct reduction effective Hamiltonian involves the additional approximation of neglecting contributions from the negative–energy intermediate states, an approximation which is unnecessary in the Foldy–Wouthuysen case as there the  $4 \times 4$  Hamiltonian does not connect positive– and negative–energy states. We conclude that at least in the cases where the relativistic Hamiltonian is known, using the direct Pauli reduction effective Hamiltonian introduces spuri-

ous relativistic effects and therefore the Foldy–Wouthuysen reduction should be preferred.

## 1 Introduction

There are many cases in low– and intermediate–energy nuclear physics where one is confronted with the problem of incorporating relativistic effects in a nonrelativistic calculation. Such a situation arises, for example, when a relativistic Hamiltonian description exists for one part of the interaction but another part can be realistically described only by a nonrelativistic Hamiltonian. Naturally, as the energies involved in these processes increase, inclusion of relativistic corrections is expected to become more and more important.

There are at least two conceivable ways to construct “effective  $2 \times 2$  Hamiltonians” in order to identify relativistic corrections to nonrelativistic Hamiltonians.

If the full relativistic Hamiltonian,  $H$ , is known, the Foldy–Wouthuysen (FW) method [1, 2, 3, 4, 5, 6, 7, 8] provides a systematic procedure to block–diagonalize the Hamiltonian order by order in  $1/M$  and hence to decouple positive– and negative–energy states to any desired order in  $1/M$ . This is achieved by an application of successive unitary (in general time–dependent) transformations on the operator  $H - i\partial/\partial t$ . This method has frequently been used in the context of electromagnetic processes in hadronic systems.

Alternatively, one can think of evaluating the matrix element of the relativistic interaction Hamiltonian between free positive–energy solutions. After reducing this matrix element to two–component form one may expand the resulting expression to any order in  $1/M$  and use it as an effective Hamiltonian  $H_I^{eff-P}$  to be evaluated between nonrelativistic Pauli wave functions. We shall refer to that procedure as the “direct Pauli reduction”.

The purpose of this paper is to address the question whether these two procedures yield identical results, and if not, which of the two should be preferred, as they both are used in the literature. For example, in electron scattering from nuclei, one evaluates matrix elements of electromagnetic operators (“semi-de-relativized” according to the direct Pauli reduction procedure) between nonrelativistic solutions of the Schrödinger equation with a given  $NN$  potential [9, 10, 11, 12, 13], but there also exist treatments that make use of the Foldy–Wouthuysen procedure to obtain semi–relativistic op–

erators [8, 14, 15]. Similarly, in proton–proton bremsstrahlung the strong interaction part is conventionally treated using a nonrelativistic Lippmann–Schwinger approach, whereas for the electromagnetic interaction a semi-relativistic ansatz, involving either the Foldy–Wouthuysen method or the direct Pauli reduction method is made [16, 17, 18, 19, 20].

There are basically two observations that provide the key to our investigation. The first observation is that there should be a clear distinction between Hamiltonians that are obtained starting from a relativistic one by means of unitary transformations (and which are still  $4 \times 4$  operators) and what we shall refer to as effective  $2 \times 2$  Hamiltonians (which are  $2 \times 2$  operators, constructed as the upper left–hand parts of the corresponding  $4 \times 4$  Hamiltonians). As we shall show, the direct Pauli reduction can be also viewed as an effective  $2 \times 2$  Hamiltonian obtained from a  $4 \times 4$  unitary time–independent transformation of the relativistic Hamiltonian, by eliminating the contribution of the negative–energy sector of the relativistic Hamiltonian. This distinction extends to the S–matrices also, where the  $4 \times 4$  Hamiltonians lead to the full relativistic S–matrix, whereas the effective  $2 \times 2$  Hamiltonians lead to effective S–matrices. Such effective S–matrices, which are calculated in non-relativistic scattering theory, may, depending on the interaction, show effects of the neglect of the negative–energy sector.

The second observation is that in most practical applications both these methods for obtaining effective Hamiltonians are not applied to the *total* (that is, electroweak–plus–strong) Hamiltonian but only to the electroweak part, in combination with a nonrelativistic approach to the strong interaction. Physical observables, such as S–matrix, are then calculated in the framework of old–fashioned time–ordered perturbation theory. The examples presented above, namely, electron scattering from nuclei and proton–proton bremsstrahlung are clearly of that type.

It is commonly believed that both types of effective Hamiltonians lead to identical effective S–matrices (and hence to identical observables) and thus that it makes no difference which is used in a nonrelativistic calculation. A justification for this assumption has been given in refs. [21, 22]. We discuss this assumption in quite some detail and find that *in general* it is not correct. In particular, in the cases where the full relativistic Hamiltonian is known, we show that the transformed  $4 \times 4$  Hamiltonians all lead to the same S–matrix, which is (to the given order in  $1/M$ ) identical to the relativistic S–matrix. However, the *effective* S–matrices corresponding to the Foldy–

Wouthuysen and the direct Pauli reduction, which are those appropriate for a nonrelativistic calculation, differ from each other beyond first order in perturbation theory and only the Foldy–Wouthuysen effective S–matrix reproduces the full relativistic S–matrix to that particular order in  $1/M$ . We discuss these issues in detail for the particular case of Compton scattering by a proton in ref. [23]. In the cases where the full relativistic Hamiltonian is not known we cannot argue in favour of any of the different effective S–matrices in particular. In order to get a feeling about the importance of these differences, we compare the results of a proton–proton bremsstrahlung calculation using the two different methods.

Our paper is organized as follows. In the next section we discuss general time–dependent unitary transformations. In this context we are mainly concerned with the Foldy–Wouthuysen transformation. In the third section we compare the first–order matrix elements of the effective Hamiltonians obtained with the Foldy–Wouthuysen and direct Pauli reduction procedures and investigate under which circumstances they lead to the same result. In the fourth section we argue about the need to go beyond first–order perturbation theory. Using a generic Hamiltonian with the sole assumption of harmonic time dependence we explain how differences arise in the different effective S–matrices. In section 5 we give a general nonperturbative proof of why the S–matrix elements corresponding to the  $4 \times 4$  transformed Hamiltonians are identical with the full relativistic S–matrix. We conclude and summarize our main results in section 6.

## 2 The Foldy–Wouthuysen Transformation

In this section we discuss time–dependent unitary transformations with special emphasis on the Foldy–Wouthuysen transformation.

## 2.1 Unitary Transformations of the Schrödinger Equation

If we start with an equation of motion of the Schrödinger type<sup>1</sup>,

$$i\frac{\partial|\Psi(t)\rangle}{\partial t} = H(t)|\Psi(t)\rangle, \quad (1)$$

where we allow for an explicit time dependence of the Hamiltonian operator  $H(t)$ , a unitary transformation  $T(t)$

$$\begin{aligned} |\Psi'(t)\rangle &= T(t)|\Psi(t)\rangle, \\ T(t)T^\dagger(t) &= T^\dagger(t)T(t) = 1, \end{aligned} \quad (2)$$

will result in the new Schrödinger equation

$$i\frac{\partial|\Psi'(t)\rangle}{\partial t} = \left( T(t)H(t)T^\dagger(t) - iT(t)\frac{\partial T^\dagger(t)}{\partial t} \right) |\Psi'(t)\rangle \equiv H'(t)|\Psi'(t)\rangle. \quad (3)$$

In many applications, such as e. g. the Foldy–Wouthuysen transformation [1, 2, 3], it is useful to parameterize  $T(t)$  as

$$T(t) = e^{iS(t)}, \quad S(t) = S^\dagger(t), \quad (4)$$

and then to expand the new Hamiltonian  $H'(t)$  in terms of  $S(t)$ ,

$$\begin{aligned} H'(t) &= H + i[S, H] - \frac{1}{2}[S, [S, H]] - \frac{i}{6}[S, [S, [S, H]]] + \dots \\ &\quad - \dot{S} - \frac{i}{2}[S, \dot{S}] + \frac{1}{6}[S, [S, \dot{S}]] + \dots, \end{aligned} \quad (5)$$

with  $\dot{S} = \partial S / \partial t$ .

As the unitary transformation  $T(t)$  may in general be time-dependent, one clearly finds [3, 21, 22]

$$\langle \Psi(t) | H(t) | \Psi(t) \rangle \neq \langle \Psi'(t) | H'(t) | \Psi'(t) \rangle, \quad (6)$$

which simply expresses the fact that if  $H(t)$  is the operator corresponding to a physical observable in the first representation,  $H'(t)$  is *not* the corresponding

---

<sup>1</sup>We regard the Dirac equation as a specific example of a Schrödinger type equation.

operator describing the *same* observable in the second representation [21, 22]. In other words, as was pointed out by Nieto [21], it is the unitary transform of  $H(t)$ , i. e.  $T(t)H(t)T^\dagger(t)$ , which is physically equivalent to  $H(t)$ , as it is this operator which yields the same matrix element between transformed states as  $H(t)$  between the original states.

## 2.2 The Foldy–Wouthuysen Method

The Foldy–Wouthuysen transformation [1, 2, 3] provides a systematic method of finding a representation of the Dirac Hamiltonian in which positive– and negative–energy states are separately represented by two–component wave functions instead of the four–component wave functions in the ordinary Dirac representation. For the free Dirac equation<sup>2</sup>

$$i\frac{\partial\Psi_0(x)}{\partial t} = (\vec{\alpha} \cdot \vec{p} + \beta M)\Psi_0(x) = H_0\Psi_0(x), \quad (7)$$

where  $\alpha_i$  and  $\beta$  are the usual Dirac matrices,  $\vec{p}$  is the momentum operator, and  $x$  is a shorthand notation for  $(\vec{x}, t)$ , the transformation is exactly known and given by

$$\begin{aligned} \Psi_0^{FW}(x) &= T_0\Psi_0(x), \\ T_0 &= \sqrt{\frac{E+M}{2E}} \begin{pmatrix} 1 & \frac{\vec{\sigma}\cdot\vec{p}}{E+M} \\ -\frac{\vec{\sigma}\cdot\vec{p}}{E+M} & 1 \end{pmatrix}, \end{aligned} \quad (8)$$

where we have defined the *operator*  $E = \sqrt{\vec{p}^2 + M^2}$ . In the new Foldy–Wouthuysen representation the free Dirac equation has the simple form

$$i\frac{\partial\Psi_0^{FW}(x)}{\partial t} = \beta\sqrt{\vec{p}^2 + M^2}\Psi_0^{FW}(x) = H_0^{FW}\Psi_0^{FW}(x). \quad (9)$$

As  $\beta$  is of block–diagonal form, eq. (9) is just the direct sum of two Hamiltonians  $\pm\sqrt{\vec{p}^2 + M^2}$ . The positive/negative–energy solutions of eq. (9) are of the form

$$\begin{aligned} \Psi_0^{FW(+)}(x) &= \begin{pmatrix} \chi^{(+)}(x) \\ 0 \end{pmatrix}, \\ \Psi_0^{FW(-)}(x) &= \begin{pmatrix} 0 \\ \chi^{(-)}(x) \end{pmatrix}, \end{aligned} \quad (10)$$

---

<sup>2</sup>States and Hamiltonians without specific superscript labels will always be assumed to be in the Dirac representation.

where  $\chi^\pm(x)$  and 0 are two-component spinors. Note that  $H_0^{FW}$  is simply given by  $T_0 H_0 T_0^\dagger$  because  $T_0$  is time-independent.

If the Dirac Hamiltonian contains an explicitly time-dependent interaction  $H_I(t)$ , the transformation  $T$  will depend on time,  $T = T(t)$ , and, in general, a closed form for the transformation  $T(t)$  leading to a block-diagonal form of the Hamiltonian is not known. However, Foldy and Wouthuysen [1] developed a systematic procedure to construct a new Hamiltonian which is block-diagonal to any desired order in  $1/M$ . The idea is to split the Dirac Hamiltonian into its odd and even components<sup>3</sup>,

$$H(t) = H_0 + H_I(t) = \beta M + \mathcal{O}(t) + \mathcal{E}_I(t), \quad (11)$$

where  $\mathcal{O}(t) = \mathcal{O}_0 + \mathcal{O}_I(t) = \vec{\alpha} \cdot \vec{p} + \mathcal{O}_I(t)$ . It is then assumed, that the interaction potentials  $\mathcal{O}_I$  and  $\mathcal{E}_I$  do not contain powers of  $1/M$  that are smaller than  $(1/M)^0$ . Furthermore, the interaction has to be *weak* enough in the sense that the magnitude of each of the time and space Fourier components of the interaction potential is considerably smaller than the mass of the nucleon [1]. It is then understood that we mean by  $[1/M]$  “terms of order  $E_{ref}/M$ ”, where  $E_{ref}$  is some reference energy smaller than  $M$ . The procedure consists of first applying the transformation

$$\begin{aligned} T^{(1)}(t) &= e^{iS^{(1)}(t)}, \\ S^{(1)}(t) &= -i\beta \frac{\mathcal{O}(t)}{2M}, \end{aligned} \quad (12)$$

to eq. (11), the result of which is then written as

$$H^{FW(1)}(t) = \beta M + \mathcal{O}^{(1)}(t) + \mathcal{E}^{(1)}(t). \quad (13)$$

Using eq. (5) it can be easily shown [1, 2, 3] that the transformation of eq. (12) is constructed such that the odd operator of eq. (13) is of order  $1/M$ . The procedure is then repeated with a new transformation  $T^{(2)}(t)$  which is exactly of the same form as eq. (12) except that in  $S^{(2)}(t)$  the new odd operator  $\mathcal{O}^{(1)}(t)$  of eq. (13) appears. As each successive application reduces the leading power of the odd operator of the resulting Hamiltonian by one

---

<sup>3</sup>Odd operators  $\mathcal{O}$ , such as e. g.  $\alpha_i$ , couple large and small components whereas even operators  $\mathcal{E}$  (e. g.  $\beta$ ) do not. The following identities are useful in the derivation of the Foldy–Wouthuysen transformation:  $[\mathcal{O}, \beta] = 2\mathcal{O}\beta$ ,  $[\mathcal{E}, \beta] = 0$ .

unit, after applying this method  $n$  times one obtains a Hamiltonian which is block-diagonal to order  $1/M^{n-1}$ . For example, after four transformations the Hamiltonian which does not contain any odd operators up to and including order  $1/M^3$  reads

$$H^{FW(4)}(t) = \beta \left( M + \frac{\mathcal{O}^2}{2M} - \frac{\mathcal{O}^4}{8M^3} \right) + \mathcal{E}_I - \frac{1}{8M^2} [\mathcal{O}, [\mathcal{O}, \mathcal{E}_I] + i\dot{\mathcal{O}}_I] - \frac{\beta}{8M^3} ([\mathcal{O}, \mathcal{E}_I] + i\dot{\mathcal{O}}_I)^2 + [\frac{1}{M^4}], \quad (14)$$

where  $\dot{\mathcal{O}}_I$  stands for  $\partial \mathcal{O}_I / \partial t$ .

### 3 First-Order Matrix Elements

It is in general not possible to explicitly solve the Dirac equation with an arbitrary time-dependent potential. On the other hand in many cases it is sufficient to construct a nonrelativistic approximation including lowest-order relativistic corrections. Such an approximation is often formulated in terms of a  $2 \times 2$  nonrelativistic effective Hamiltonian  $H^{eff}$  which is to be used in nonrelativistic perturbation theory. One way of constructing such an effective Hamiltonian to be used in low- and intermediate-energy applications is to evaluate the matrix element of the potential operator between positive-energy solutions of the *free* Dirac equation, perform a two-component reduction and then interpret the result as the matrix element of an effective  $2 \times 2$  Hamiltonian between nonrelativistic (positive-energy) states (direct Pauli reduction). Commonly the so constructed effective Hamiltonian is expanded in a series in  $1/M$  and only the terms to a particular order are kept.

#### 3.1 Reduction of the First-Order Matrix Element

To be specific we consider the expression

$$\int d^3x \Psi_{0f}^\dagger(\vec{x}, t) H_I(t) \Psi_{0i}(\vec{x}, t) = \int d^3x \Psi_{0f}^{FW\dagger}(\vec{x}, t) T_0 H_I(t) T_0^\dagger \Psi_{0i}^{FW}(\vec{x}, t), \quad (15)$$

where  $\Psi_{0i}$  and  $\Psi_{0f}$  are *positive-energy* solutions of the free Dirac equation and  $\Psi_{0i}^{FW}$  and  $\Psi_{0f}^{FW}$  are the corresponding solutions of the Foldy-Wouthuysen Hamiltonian (see eq. (8) and (9)). Expressions of the type of eq. (15) typically



appear as building blocks of a perturbative treatment of the S-matrix. From eq. (15) we find that instead of explicitly reducing the matrix element, we may as well look at the operator  $H_I^P = T_0 H_I(t) T_0^\dagger$ . This operator, when taken between the states  $\Psi_{0i}^{FW}$  and  $\Psi_{0f}^{FW}$ , leads to an effective  $2 \times 2$  Hamiltonian  $H_I^{eff-P}$  which is just what one would get by performing a two-component Pauli reduction on the left-hand side of eq. (15).

For a general interaction,  $H_I(t) = \mathcal{O}_I(t) + \mathcal{E}_I(t)$ , a reduction through order  $1/M^3$  yields

$$\begin{aligned} H_I^P(t) = & \mathcal{E}_I + \frac{\beta}{2M} \{\mathcal{O}_0, \mathcal{O}_I\} - \frac{1}{8M^2} [\mathcal{O}_0, [\mathcal{O}_0, \mathcal{E}_I]] - \frac{\beta}{8M^3} \{\mathcal{O}_0^2, \{\mathcal{O}_0, \mathcal{O}_I\}\} \\ & - \frac{\beta}{16M^3} [\mathcal{O}_0^2, [\mathcal{O}_0, \mathcal{O}_I]] + [1/M^4] + \text{odd terms.} \end{aligned} \quad (16)$$

Here the terms written explicitly are all even operators. The odd terms do not contribute when eq. (16) is evaluated between free positive-energy states of the Foldy-Wouthuysen representation. However, as it is only these states which are considered in a nonrelativistic calculation, one can regard the upper left-hand block of  $H_I^P(t)$  as an effective Hamiltonian  $H_I^{eff-P}(t)$  to be used with two-component Pauli spinors.

### 3.2 Comparison with the Foldy-Wouthuysen Hamiltonian

The Foldy-Wouthuysen Hamiltonian of eq. (14) may be written as

$$H^{FW(4)} = H_0^{FW} + H_1^{FW}(t) + H_2^{FW}(t) + H_3^{FW}(t) + H_4^{FW}(t) + [1/M^4], \quad (17)$$

where the subscripts  $i$  indicate the order in which the interaction potentials  $\mathcal{O}_I$  and  $\mathcal{E}_I$  appear in the corresponding Hamiltonian. The *linear* interaction Hamiltonian is given by

$$\begin{aligned} H_1^{FW}(t) = & \mathcal{E}_I + \frac{\beta}{2M} \{\mathcal{O}_0, \mathcal{O}_I\} - \frac{1}{8M^2} [\mathcal{O}_0, [\mathcal{O}_0, \mathcal{E}_I] + i\dot{\mathcal{O}}_I] \\ & - \frac{\beta}{8M^3} \{\mathcal{O}_0^2, \{\mathcal{O}_0, \mathcal{O}_I\}\} + [1/M^4]. \end{aligned} \quad (18)$$

Comparing eq. (16) and eq. (18) we see that the two operators  $H_1^{FW}(t)$  and  $H_I^P(t)$  differ by

$$\Delta H_1(t) = H_1^{FW}(t) - H_I^P(t)$$

$$\begin{aligned}
&= -\frac{i}{8M^2}[\mathcal{O}_0, \dot{\mathcal{O}}_I] + \frac{\beta}{16M^3}[\mathcal{O}_0^2, [\mathcal{O}_0, \mathcal{O}_I]] + \text{odd terms} + [1/M^4] \\
&= \frac{1}{8M^2}([H_0^{FW}, [\mathcal{O}_0, \mathcal{O}_I]] - i[\mathcal{O}_0, \dot{\mathcal{O}}_I]) + \text{odd terms} + [1/M^4].
\end{aligned}
\tag{19}$$

In obtaining the last line of eq. (19) we expanded eq. (9) and made use of  $\mathcal{O}_0^2 = \vec{p}^2$ .

It has often (tacitly) been assumed that  $\Delta H_1(t)$  of eq. (19) vanishes when evaluated between free positive-energy Foldy–Wouthuysen states. In which case, in a nonrelativistic calculation to first order in the interaction, it should make no difference which of the two Hamiltonians  $H_I^P(t)$  and  $H_I^{FW}(t)$  is used for the interaction. We will now investigate under which circumstances this assumption is correct.

Most applications of the Foldy–Wouthuysen transformation are concerned with either static (atomic physics, hydrogen atom) or harmonically time-dependent external potentials. If we assume the potential  $\mathcal{O}_I(t)$  to have a harmonic time dependence, we find for the time derivative  $\dot{\mathcal{O}}_I = \mp i\omega\mathcal{O}_I$ , where the upper/lower sign refers to the absorption/emission of a quantum of energy  $\omega$ . In this case the matrix element of  $\Delta H_1(t)$  reads

$$\begin{aligned}
&\int d^3x \Psi_{0f}^{FW\dagger}(\vec{x}, t) \Delta H_1(t) \Psi_{0i}^{FW}(\vec{x}, t) = \\
&\frac{1}{8M^2}(E_f - E_i \mp \omega) \int d^3x \Psi_{0f}^{FW\dagger}(\vec{x}, t) [\mathcal{O}_0, \mathcal{O}_I] \Psi_{0i}^{FW}(\vec{x}, t) + [1/M^4].
\end{aligned}
\tag{20}$$

From eq. (20) it is seen, that for a potential with harmonic time dependence, both Hamiltonians will yield the same first-order matrix elements if the energies of the initial and final state match with the absorption/emission of a quantum of energy  $\omega$ , or to put it in a somewhat sloppy way, if energy is “conserved” in the transition, or alternatively if the commutator  $[\mathcal{O}_0, \mathcal{O}_I]$  vanishes, which in general we do not expect to happen. This observation has also been made in ref. [7].

In second-order (old-fashioned) perturbation theory, energy is not conserved at a vertex [26, 27]. This clearly suggests, that in practical second-order calculations, such as e. g. proton–proton bremsstrahlung, it may make a difference whether one uses the linear effective Foldy–Wouthuysen Hamil-

tonian  $H_1^{eff-FW}(t)$  deduced from eq. (18) or the Hamiltonian  $H_I^{eff-P}(t)$  obtained from eq. (16).

Our result seems to be somewhat different from the earlier work of Nieto [21] and of Goldman [22], who appear to have shown that the results from the two first-order Hamiltonians are the same, i. e. that the matrix element of  $\Delta H_1(t) = 0$  in general. To understand this difference, consider again this matrix element. Using the last part of eq. (19) one can write

$$\begin{aligned} \int d^3x \Psi_{0f}^{FW\dagger}(\vec{x}, t) \Delta H_1(t) \Psi_{0i}^{FW}(\vec{x}, t) = \\ - \int d^3x \Psi_{0f}^{FW\dagger}(\vec{x}, t) \left( [\Delta, H_0^{FW}] + i\dot{\Delta} \right) \Psi_{0i}^{FW}(\vec{x}, t) + [1/M^4], \end{aligned} \quad (21)$$

where to the order being considered  $\Delta = [\mathcal{O}_0, \mathcal{O}_I]/8M^2$ . A more general form for  $\Delta$  is given in ref. [22], but is unnecessary for the purposes of this discussion.

The authors of both ref. [21] and ref. [22] make the *assumption* that

$$[\Delta, H_0^{FW}] + i\dot{\Delta} = 0, \quad (22)$$

which they refer to as the ‘‘Schrödinger statement’’, and thus conclude that  $\langle \Psi_{0f}^{FW\dagger}(t) | \Delta H_1(t) | \Psi_{0i}^{FW}(t) \rangle \equiv 0$  and consequently that it makes no difference which of the Hamiltonians is used.

It is clear however from eq. (19) that eq. (22) cannot be correct in general as an operator equation. If one considers it instead as a matrix equation, then the Schrödinger equation can be used to replace the Hamiltonian by a time derivative and one obtains

$$\begin{aligned} \langle \Psi_{0f}^{FW}(t) | \left( [\Delta, H_0^{FW}] + i\dot{\Delta} \right) | \Psi_{0i}^{FW}(t) \rangle = \\ i \frac{d}{dt} \langle \Psi_{0f}^{FW}(t) | \Delta | \Psi_{0i}^{FW}(t) \rangle + [1/M^4]. \end{aligned} \quad (23)$$

Again, however, the right-hand side of this equation is not identically zero in general. If one assumes a harmonic time dependence of the interaction, as we have done above, then the time dependence of the matrix element is just  $e^{i(E_f - E_i \mp \omega)t}$  and the time derivative simply brings down the energy factor  $i(E_f - E_i \mp \omega)$ . This leads exactly to our previous eq. (20) and to our previous conclusion that the matrix element of  $\Delta H_1$  vanishes, and the results using different Hamiltonians are equivalent, only when the transition

conserves energy. In the more general case when the states are off-energy-shell, as will be the case when such matrix elements are used in second-order perturbation theory, different Hamiltonians may be expected to give different results.

As a specific example of a case where the choice of Hamiltonian does make a difference, let us consider the interaction of a proton with an external electromagnetic field,

$$\begin{aligned}\mathcal{E}_I(t) &= e\Phi(\vec{x}, t) - \frac{e\kappa}{2M}\beta\vec{\sigma} \cdot \vec{B}(\vec{x}, t), \\ \mathcal{O}_I(t) &= -e\vec{\alpha} \cdot \vec{A}(\vec{x}, t) + i\frac{e\kappa}{2M}\beta\vec{\alpha} \cdot \vec{E}(\vec{x}, t),\end{aligned}\quad (24)$$

where  $e > 0$  is the elementary charge,  $\kappa = 1.79$  the anomalous magnetic moment in units of a nuclear magneton, and  $\vec{E} = -\vec{\nabla}\Phi - \dot{\vec{A}}$  and  $\vec{B} = \vec{\nabla} \times \vec{A}$ . Using this specific example the *even* part of  $\Delta H_1(t)$  is given to order  $1/M^3$  by

$$\begin{aligned}\Delta H_1(t) &= -\frac{e\beta}{16M^3}[\vec{p}^2, -i\vec{\nabla} \cdot \vec{A} + i\vec{\sigma} \cdot \vec{p} \times \vec{A} - i\vec{\sigma} \cdot \vec{A} \times \vec{p}] \\ &\quad -\frac{e\kappa}{16M^3}\beta(\vec{p} \cdot \dot{\vec{E}} + \dot{\vec{E}} \cdot \vec{p} + \vec{\sigma} \cdot \vec{\nabla} \times \dot{\vec{E}}) \\ &\quad +\frac{ie}{8M^2}(-i\vec{\nabla} \cdot \dot{\vec{A}} + i\vec{\sigma} \cdot \vec{p} \times \dot{\vec{A}} - i\vec{\sigma} \cdot \dot{\vec{A}} \times \vec{p}).\end{aligned}\quad (25)$$

In all modern proton-proton bremsstrahlung calculations [16, 17, 18, 19, 20] the electromagnetic part of the interaction is obtained by reducing the relativistic nucleon-nucleon-gamma vertex to obtain an effective  $2 \times 2$  Hamiltonian which is used with two-component spinors in a nonrelativistic calculation. In some cases, e. g. ref. [16], the effective Hamiltonian  $H_1^{eff-FW}$  obtained via a Foldy-Wouthuysen reduction as in eq. (18) was used and in others, e. g. [17], the  $H_I^{eff-P}$  obtained via a direct two-component reduction as in eq. (16) was used. These two should differ by the  $\Delta H_1$  of eq. (19) or eq. (25).

In figure 1 the results of such a bremsstrahlung calculation are shown for a particular kinematic situation, corresponding to a coplanar equal-angle geometry with outgoing protons measured at  $\Theta_3 = \Theta_4 = 6^\circ$  and with an incident laboratory energy  $T^{LAB} = 280$  MeV. The calculations of ref. [16] have been used taking a) the Foldy-Wouthuysen Hamiltonian for the electromagnetic interaction of the proton or b) the effective Hamiltonian of eq. (16). As

can be seen from eq. (20) the difference between the two calculations is of order  $1/M^2$  multiplied by the amount by which energy is not conserved at the electromagnetic vertex. For the kinematics chosen in figure 1 the difference is of order 7 %.

We conclude that as long as one considers only first-order matrix elements, such as those appearing in  $\mu + p \rightarrow n + \nu$  or electron scattering from an on-shell nucleon [24] it does not make a difference which reduction scheme is used. This is true because in a perturbative treatment of the S-matrix the time integration enforces energy conservation at the vertex of lowest-order perturbation theory. However, as we have seen in figure 1 the situation is different for higher-order processes such as e. g. proton-proton bremsstrahlung.

One might properly fault this particular numerical example because the strong interaction is not treated consistently with the electromagnetic interaction, i. e. not as a reduction of some relativistic interaction. Thus it is important to look at more detail at second-order processes in which both interactions can be treated in the same way. That we do in general in the following section and for the specific case of Compton scattering by a proton in ref. [23].

## 4 Second-Order Perturbation Theory

In the previous sections we described how to obtain various effective interaction Hamiltonians for use in nonrelativistic calculations and showed that the first-order matrix elements of these Hamiltonians differed only in situations when the states involved did not conserve energy. Such a situation arises in second- (or higher-) order old-fashioned time-ordered perturbation theory where such first-order matrix elements connect to an intermediate state. Of course, the *total* energy is conserved for each (higher-order) diagram, but energy is not conserved at the individual vertices of such a diagram [26, 27]. It is this fact which *may* lead to different predictions for the effective S-matrix and thus it is important to see what happens in such second-order processes.

There are two questions which have to be addressed. Firstly, which kind of transformation of the  $4 \times 4$  Hamiltonian leads to the same relativistic S-matrix? As long as we are dealing with Hamiltonians originating from a unitary transformation of the states (see section 2.1) one would expect that

the full relativistic S-matrix would be invariant under such transformations. That is indeed the case as we will show in second-order perturbation theory in the first part of this section and by more formal manipulations valid to all orders in sec. 5.

This invariance of the full relativistic S-matrix has been observed by others, and is consistent with our usual understanding of quantum mechanics. However it has led to the extremely misleading, and incorrect, assumption that the physical observables calculated with various *effective* Hamiltonians will be the same and thus that it makes no difference which effective Hamiltonian is used.

Hence we are led to the second question which is of great practical importance. Do different effective  $2 \times 2$  Hamiltonians lead to the same non-relativistic effective S-matrix? We show in the last part of this section that in fact this is not the case. The reason is that the construction of effective Hamiltonians is not always done solely via a unitary transformation of the states. Thus it *does* make a difference which effective Hamiltonian is used in nonrelativistic calculations. Only the Foldy-Wouthuysen Hamiltonian (and other block-diagonal Hamiltonians obtained from it) reproduces the full relativistic S-matrix to a given order in  $1/M$ . Thus it is only such Hamiltonians which can be presumed to give correct physical results in nonrelativistic calculations.

## 4.1 Full Relativistic S-Matrix in Second-Order Perturbation Theory

We now restrict our consideration to second-order perturbation theory. Thus we assume that the relativistic  $4 \times 4$  Dirac Hamiltonian is given as before by  $H = H_0 + H_I(t)$ , where  $H_0$  is the free Dirac Hamiltonian. To be definite we will focus on processes involving an incoming quantum of energy  $\omega_a$  and an outgoing one of energy  $\omega_b$ . Thus  $H_I(t)$  will have a term  $\sim H_a e^{-i\omega_a t}$  and one  $\sim H_b e^{i\omega_b t}$  as well as others as required for Hermiticity. Formally we could construct  $H_I(t)$  in second quantized form and use appropriate creation and destruction operators to pick out the pieces needed. Instead we will treat  $H_I(t)$  somewhat loosely with the understanding that, at the end, by second-order we will mean that we will discard all terms except those containing one power of the coupling associated with  $H_a$  and one power of that associated

with  $H_b$ . It may be that  $H_a = H_b$ , as for example in Compton scattering where  $H_a = H_b \sim H_{\gamma NN}$ . Alternatively  $H_a$  may be different from  $H_b$  as in pion photoproduction where we could take  $H_a \sim H_{\gamma NN}$  and  $H_b \sim H_{\pi NN}$ .

With these assumptions the full relativistic S-matrix is given through second order in the interaction in the standard fashion, e. g. ref. [2], eq. (6.57), as

$$S_{fi} = \delta_{fi} - i \int d^4y \Psi_{0f}^\dagger(y) H_I(y) \Psi_{0i}(y) - i \int d^4y d^4z \Psi_{0f}^\dagger(y) H_I(y) S_F(y-z) \beta H_I(z) \Psi_{0i}(z), \quad (26)$$

where  $\Psi_{0f}$  and  $\Psi_{0i}$  are *positive*-energy eigenstates of the free Dirac equation and where  $S_F$  is the free Feynman propagator. To proceed further we need to examine how the states, Hamiltonian, and propagator transform under the unitary transformation  $T(t)$ .

We define the states  $\Psi_{0p}(x)$  to be eigenstates of the free Dirac Hamiltonian  $H_0$ . They may have positive or negative energy and can be written as

$$\Psi_{0p}^{(\pm)}(\vec{x}, t) = e^{\mp i E_p t} \Phi_{0p}^{(\pm)}(\vec{x}), \quad (27)$$

where  $E_p = +\sqrt{\vec{p}^2 + m^2}$  and

$$\Phi_{0p}^{(\pm)}(\vec{x}) = w^{(\pm)}(p) \frac{e^{\pm i \vec{p} \cdot \vec{x}}}{(2\pi)^{\frac{3}{2}}}, \quad (28)$$

where  $w^{(+)} = u$  and  $w^{(-)} = v$  with  $u^\dagger u = v^\dagger v = 1$  are the usual Dirac spinors<sup>4</sup>. Spins will always be summed, so all explicit spin dependence will be suppressed. These states satisfy a completeness relation

$$\sum_{\text{spins}} \int d^3p \left\{ \Phi_{0p}^{(+)}(\vec{x}) \Phi_{0p}^{(+)\dagger}(\vec{y}) + \Phi_{0p}^{(-)}(\vec{x}) \Phi_{0p}^{(-)\dagger}(\vec{y}) \right\} = \delta^3(\vec{x} - \vec{y}). \quad (29)$$

Under the unitary transformation  $T_0$ , which is all that will be needed, we have  $\Psi_{0p}^{FW(\pm)}(x) = T_0 \Psi_{0p}^{(\pm)}(x)$ . This changes only the spinors which become, using the specific  $T_0$  of eq. (8),

$$u(p) \rightarrow \begin{pmatrix} \chi \\ 0 \end{pmatrix} \quad \text{and} \quad v(p) \rightarrow \begin{pmatrix} 0 \\ \chi \end{pmatrix}, \quad (30)$$

---

<sup>4</sup>Note that we use a different normalization convention in comparison with ref. [2].

with  $\chi$  a two-component spinor and  $0$  a two-component null spinor. We will always assume that  $T_0$  is obtained from  $T(t)$  by turning off the interaction, which thus removes the time dependence. For most of this section however it is not necessary to specify  $T_0$  or  $T(t)$  explicitly. Thus we will use  $\Psi'$  and  $H'$  for example for states and Hamiltonians obtained via general transformations as in sec. 2.1. We will reserve  $\Psi^{FW}$  and  $H^{FW}$  for situations where it is important that the states have the form eq. (30) and that the Hamiltonian is block-diagonal to some order in  $1/M$ , and thus where we require the specific transformation  $T_0$  of eq. (8) and the specific  $T(t)$  of sec. 2.2.

To determine how the propagator  $S_F$  transforms it is easiest to first expand in the complete set of states just defined. Thus starting with, e. g. eq. (6.48) of ref. [2], and using a standard expansion for  $\theta(t - t')$  we obtain

$$S_F(y - z)\beta = \frac{1}{2\pi} \int dp_0 d^3p e^{-ip_0(t_y - t_z)} \sum_{\text{spins}} \left\{ e^{-iE_p(t_y - t_z)} \frac{\Phi_{0p}^{(+)}(\vec{y})\Phi_{0p}^{(+)\dagger}(\vec{z})}{p_0 + i\epsilon} - e^{iE_p(t_y - t_z)} \frac{\Phi_{0p}^{(-)}(\vec{y})\Phi_{0p}^{(-)\dagger}(\vec{z})}{-p_0 + i\epsilon} \right\}. \quad (31)$$

We now define

$$S'_F(y - z)\beta = T_0 S_F(y - z)\beta T_0^\dagger. \quad (32)$$

It is clear by applying  $T_0 \dots T_0^\dagger$  to the right-hand side of eq. (31) that  $S'_F(x - y)\beta$  has exactly the same form as  $S_F(x - y)\beta$ , only with the states  $\Phi$  replaced by  $\Phi'$  and so is the correct free Feynman propagator in the  $\Phi'$  representation.

Now consider the Hamiltonian. To allow an expansion to second order we write the transformation variously as

$$T(t) = e^{iS} = T_0 e^{i\tilde{S}} = T_0 e^{i(S_1 + S_2 + \dots)}, \quad (33)$$

where  $S_n$  is Hermitian and of order  $n$  in the interaction. Starting from  $H' = T(t)H(t)T^\dagger(t) - iT(t)\dot{T}^\dagger(t)$  and inverting, we find for  $H$

$$H = e^{-i\tilde{S}} \left\{ T_0^\dagger H' T_0 + i e^{i\tilde{S}} \frac{\partial e^{-i\tilde{S}}}{\partial t} \right\} e^{i\tilde{S}}. \quad (34)$$



We now write  $H'$  as  $H' = H'_0 + H'_1 + H'_2 + \dots$ , where  $H'_n$  is of order  $n$  in the interaction and expand  $H$  in a fashion analogous to eq. (5) to obtain through second order

$$H_I = \{T_0^\dagger H'_1 T_0 - \Delta H'_1\} + \{T_0^\dagger H'_2 T_0 - \Delta H'_2\}, \quad (35)$$

$$\Delta H'_1 = T_0^\dagger \Delta H_1 T_0 = i[S_1, H_0] - \dot{S}_1, \quad (36)$$

$$\Delta H'_2 = i[S_2, H_0] - \dot{S}_2 - \frac{i}{2}[S_1, \Delta H'_1] + i[S_1, T_0^\dagger H'_1 T_0]. \quad (37)$$

$H'_2$  is second-order in the interaction and is a contact term which is generated by the transformation.  $H_I$  is however first-order in the interaction, so the second bracketed expression in eq. (35) must be zero. Hence  $\Delta H'_2 = T_0^\dagger H'_2 T_0$ .

The next step is to substitute this result for  $H_I$  into eq. (26) and use eq. (32) for  $S'_F(x-y)\beta$  to obtain a result for that part of the full relativistic S-matrix which is second-order in the interaction,

$$\begin{aligned} S_{fi} = & -i \int d^4 y \Psi_{0f}^\dagger(y) H'_2(y) \Psi'_{0i}(y) \\ & -i \int d^4 y d^4 z \Psi_{0f}^\dagger(y) H'_1(y) S'_F(y-z) \beta H'_1(z) \Psi'_{0i}(z) \\ & + \Delta S_{fi}, \end{aligned} \quad (38)$$

where

$$\begin{aligned} \Delta S_{fi} = & i \int d^4 y \Psi_{0f}^\dagger(y) \Delta H'_2(y) \Psi_{0i}(y) \\ & -i \int d^4 y d^4 z \Psi_{0f}^\dagger(y) \left\{ -T_0^\dagger H'_1(y) T_0 S_F(y-z) \beta \Delta H'_1(z) \right. \\ & - \Delta H'_1(y) S_F(y-z) \beta T_0^\dagger H'_1(y) T_0 \\ & \left. + \Delta H'_1(y) S_F(y-z) \Delta H'_1(z) \right\} \Psi_{0i}(z), \end{aligned} \quad (39)$$

Observe now that if  $\Delta S_{fi} = 0$ , as we will in fact show, then the right-hand side of eq. (38) is exactly the S-matrix that one would compute to second-order perturbation theory using the transformed Hamiltonian  $H'$  and the transformed states. Comparison with eq. (26) shows then that the S-matrix is invariant under the unitary transformations of the type we have considered.

Thus consider  $\Delta S_{fi}$ . It is sufficient to look only at the  $\Delta H \times \Delta H$  term as the others are all similar. This term becomes after using the explicit

expression for  $S_F$ , eq. (31), and extracting the time dependence from the interaction and the states

$$\begin{aligned}
& -\frac{i}{2\pi} \int dt_y dt_z dp_0 d^3p \times \\
& \left[ \sum_{\text{spins}} \left\{ e^{-iA_+ t_y + iB_+ t_z} \frac{\langle \Phi_{0f} | \Delta H'_{1b} | \Phi_{0p}^{(+)} \rangle \langle \Phi_{0p}^{(+)} | \Delta H'_{1a} | \Phi_{0i} \rangle}{p_0 + i\epsilon} \right. \right. \\
& \quad \left. \left. - e^{-iA_- t_y + iB_- t_z} \frac{\langle \Phi_{0f} | \Delta H'_{1b} | \Phi_{0p}^{(-)} \rangle \langle \Phi_{0p}^{(-)} | \Delta H'_{1a} | \Phi_{0i} \rangle}{-p_0 + i\epsilon} \right\} + \text{c. t.} \right] \\
& = -2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \int d^3p \times \\
& \left[ \sum_{\text{spins}} \left\{ \frac{\langle \Phi_{0f} | \Delta H'_{1b} | \Phi_{0p}^{(+)} \rangle \langle \Phi_{0p}^{(+)} | \Delta H'_{1a} | \Phi_{0i} \rangle}{E_f + \omega_b - E_p + i\epsilon} \right. \right. \\
& \quad \left. \left. + \frac{\langle \Phi_{0f} | \Delta H'_{1b} | \Phi_{0p}^{(-)} \rangle \langle \Phi_{0p}^{(-)} | \Delta H'_{1a} | \Phi_{0i} \rangle}{E_f + \omega_b + E_p - i\epsilon} \right\} + \text{c. t.} \right], \quad (40)
\end{aligned}$$

where  $A_{\pm} = p_0 - E_f - \omega_b \pm E_p$  and  $B_{\pm} = p_0 - E_i - \omega_a \pm E_p$ . Here c. t. indicates that there is a standard cross term which has not been written explicitly.

Using the time dependence of  $S_1$ , which is linear in  $H_1$ , the matrix elements of  $\Delta H'_1$  can be evaluated as

$$\begin{aligned}
\langle \Phi_{0p}^{(\pm)} | \Delta H'_{1a} | \Phi_{0i} \rangle & = \langle \Phi_{0p}^{(\pm)} | i[S_{1a}, H_0] + i\omega_a S_{1a} | \Phi_{0i} \rangle \\
& = i(E_i + \omega_a \mp E_p) \langle \Phi_{0p}^{(\pm)} | S_{1a} | \Phi_{0i} \rangle, \quad (41)
\end{aligned}$$

$$\begin{aligned}
\langle \Phi_{0f} | \Delta H'_{1b} | \Phi_{0p}^{(\pm)} \rangle & = \langle \Phi_{0f} | i[S_{1b}, H_0] - i\omega_b S_{1b} | \Phi_{0p}^{(\pm)} \rangle \\
& = -i(E_f + \omega_b \mp E_p) \langle \Phi_{0f} | S_{1b} | \Phi_{0p}^{(\pm)} \rangle. \quad (42)
\end{aligned}$$

These matrix elements are analogous to that of eq. (20) but, since they connect to intermediate states  $\Phi_{0p}^{(\pm)}$ , energy is not conserved and the energy factors do not vanish.

Finally, using these equations in a symmetric way (so as to end up with the commutator) and using the completeness relation eq. (29), one finds for

eq. (40)

$$-2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \langle \Phi_{0f} | -\frac{i}{2} [S_1, \Delta H'_1] | \Phi_{0i} \rangle. \quad (43)$$

To simplify the notation, the formula has been rewritten with the full interaction, thus reintroducing the  $H_a \times H_a$  and  $H_b \times H_b$  terms which should in the end be dropped.

The other terms are evaluated in exactly the same way so that one obtains

$$\begin{aligned} \Delta S_{fi} &= 2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \langle \Phi_{0f} | \Delta H'_2 + \frac{i}{2} [S_1, \Delta H'_1] \\ &\quad - i[S_1, T_0^\dagger H'_1 T_0] | \Phi_{0i} \rangle \\ &= 2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \langle \Phi_{0f} | i[S_2, H_0] - i(\omega_b - \omega_a) S_2 | \Phi_{0i} \rangle \\ &= 2\pi \delta(E_f + \omega_b - E_i - \omega_a) (E_f + \omega_b - E_i - \omega_a) \langle \Phi_{0f} | S_2 | \Phi_{0i} \rangle \\ &= 0. \end{aligned} \quad (44)$$

Here the first reduction uses the definition of  $\Delta H'_2$  of eq. (37) and the second uses overall conservation of energy.

In summary, we have in this section considered the full relativistic S-matrix in second-order perturbation theory (though to all orders in  $1/M$ ). For definiteness we focussed only on those processes involving an incoming quantum of energy  $\omega_a$  and outgoing quantum of energy  $\omega_b$ . In that context we could show explicitly that the  $\Delta S_{fi}$  of eq. (38) was zero. The remaining terms on the right-hand side of that equation give just the full relativistic S-matrix expressed in terms of the transformed Hamiltonian  $H'_I$  and the transformed states  $\Psi'_0$  to second order. Thus we have shown explicitly that the full relativistic S-matrix is unchanged under these transformations in second-order perturbation theory, which is a special case of the more general, but more formal, proof to be given below.

Note that nowhere in this proof did we use any of the properties of  $T(t)$ , except those of eq. (33). Thus the result that the full relativistic S-matrix is invariant holds for the transformation  $T(t)$  of sec. 2.2 which block-diagonalizes  $H_I$  and produces the Foldy-Wouthuysen Hamiltonian  $H_I^{FW}$ . It also holds for  $T(t) = T_0$ , where the specific  $T_0$  of eq. (8) is meant, which produces  $H_I^P$ .

One other observation is worth making. The matrix element of  $\Delta H$  is essentially the difference between the matrix elements of  $H_I$  in the  $\Phi_0$  basis

and  $H'_I$  in the  $\Phi'_0$  basis. From the details of the proof that  $\Delta S_{fi} = 0$ , as outlined in eqs. (40) - (44), one can see that these  $\Delta H$  terms cancel the energy denominators and thus generate contact terms. In general these contact terms which make up  $\Delta H_2$  came *both* from positive- and negative-energy intermediate states. This is somewhat in contrast to the generally held belief that it is the negative-energy intermediate states alone which are responsible for contact terms. It agrees however with the specific result we have obtained in ref. [23]. There the specific example of Compton scattering by a proton is described in detail and we can also see at which order in  $1/M$  the various terms arise.

## 4.2 Comparison of Effective Nonrelativistic S-matrices in Second-Order Perturbation Theory

We now want to return to the main question, namely whether or not different reductions of the relativistic interaction to effective nonrelativistic interactions lead in practical calculations to different results. To do this we need first to review how these effective Hamiltonians are obtained and used.

Usually one starts with a supposedly known relativistic interaction and makes a transformation on it followed by a projection onto the upper left-hand block which leads to a  $2 \times 2$  effective interaction Hamiltonian,  $H_I^{eff}$ , which is then used together with positive-energy two-component states to construct an effective S-matrix,  $S_{fi}^{eff}$ , according to the usual rules of non-relativistic scattering theory. Terms of higher than leading order in  $1/M$  included in the effective Hamiltonian provide relativistic corrections to the lowest order nonrelativistic result and one hopes that if enough terms are included  $S_{fi}^{eff}$  will approach the full relativistic, and thus presumed correct, result,  $S_{fi}$ .

Thus the question which is relevant for practical calculations is whether or not these effective S-matrices are the same for different reductions of the relativistic Hamiltonian, i. e. , for different  $H_I^{eff}$ , and whether they equal the full relativistic  $S_{fi}$ .

In the Foldy-Wouthuysen case the new  $H_I^{FW}$  is obtained via a transformation  $T(t)$  of the type we have been considering, constrained by the condition that to some order in  $1/M$  the  $4 \times 4$  Hamiltonian  $H_I^{FW}$  is block-

diagonal. The  $2 \times 2$  effective Hamiltonian  $H_I^{eff-FW}$  is then just the upper left-hand block, i. e. that part which has nonzero matrix element between free positive-energy states  $\Phi_0^{FW}$ . Note that by virtue of the fact that  $T(t)$  depends on the interaction,  $H_I^{eff-FW}$  will have terms of higher order in the interaction, even though  $H_I$  was first-order.

An alternative approach often used is to make a direct Pauli reduction of the relativistic matrix element  $\langle \Phi_{0f} | H_I | \Phi_{0i} \rangle = \langle \Phi_{0f}^{FW} | T_0 H_I T_0^\dagger | \Phi_{0i}^{FW} \rangle$ . Thus in this case the  $4 \times 4$  Hamiltonian  $H_I^P = T_0 H_I T_0^\dagger$  and the  $2 \times 2$   $H_I^{eff-P}$  is just the upper left-hand block of this. Note that in this case  $H_I^P$  is not block-diagonal and that  $H_I^{eff-P}$  may or may not be expanded (or truncated) in powers of  $1/M$ . Furthermore  $H_I^{eff-P}$  is linear in the interaction, so there are no contact terms generated naturally.

To see how the effective S-matrices calculated from these effective Hamiltonians are related to the full relativistic S-matrix we start with eq. (38) above with  $\Delta S_{fi} = 0$  and use eq. (31) and eq. (32) to expand the propagator into positive- and negative-energy states. The time dependence can then be extracted and the formula simplified in exactly the same way the  $\Delta H \times \Delta H$  term was treated. The result for the second-order contributions only is

$$\begin{aligned}
S_{fi} = & -2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \left\{ \langle \Phi_{0f}^{FW} | H'_2 | \Phi_{0i}^{FW} \rangle \right. \\
& + \sum_{\text{spins}} \int d^3 p \frac{\langle \Phi_{0f}^{FW} | H'_{1b} | \Phi_{0p}^{FW(+)} \rangle \langle \Phi_{0p}^{FW(+)} | H'_{1a} | \Phi_{0i}^{FW} \rangle}{E_f + \omega_b - E_p + i\epsilon} \\
& \left. + \frac{\langle \Phi_{0f}^{FW} | H'_{1b} | \Phi_{0p}^{FW(-)} \rangle \langle \Phi_{0p}^{FW(-)} | H'_{1a} | \Phi_{0i}^{FW} \rangle}{E_f + \omega_b + E_p - i\epsilon} + \text{c. t.} \right\}.
\end{aligned} \tag{45}$$

The term with  $H'_2$  and the terms with  $\Phi_{0p}^{FW(+)}$  involve only positive-energy states, which have only upper components. Hence these terms pick out just the upper left-hand block of  $H'_1$  and  $H'_2$  and consequently make up just the effective S-matrix one would calculate using the  $2 \times 2$   $H_I^{eff}$ . Thus we can write

$$\begin{aligned}
S_{fi} = & S_{fi}^{eff} - 2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \times \\
& \left\{ \sum_{\text{spins}} \int d^3 p \frac{\langle \Phi_{0f}^{FW} | H'_{1b} | \Phi_{0p}^{FW(-)} \rangle \langle \Phi_{0p}^{FW(-)} | H'_{1a} | \Phi_{0i}^{FW} \rangle}{E_f + \omega_b + E_p - i\epsilon} + \text{c. t.} \right\}.
\end{aligned}$$

(46)

Now it is obvious that for the Foldy–Wouthuysen reduction the last term vanishes, since there  $H_I^{FW}$  is block–diagonal and since a block–diagonal  $H_I^{FW}$  does not connect positive– and negative–energy states. Thus in this case  $S_{fi} = S_{fi}^{eff-FW}$  to some order in  $1/M$ . Hence the Foldy–Wouthuysen procedure gives the correct effective S–matrix when  $H_I^{eff-FW}$  is used as a  $2 \times 2$  nonrelativistic Hamiltonian in standard nonrelativistic scattering theory.

Note that the only thing needed to get this result is an interaction  $H_I^{FW}$  which is block–diagonal to a given order in  $1/M$ . Thus a further transformation on the Hamiltonian which preserves the block–diagonal nature to the same order in  $1/M$  will also give the same correct S–matrix  $S_{fi}^{eff-FW}$ . Thus it is clear that the Foldy–Wouthuysen algorithm is not unique. It has been observed by Barnhill and others (see [4] and references therein) that the order by which different, noncommuting odd operators are eliminated from the relativistic Hamiltonian may lead to such a “freedom” in the choice of the FW Hamiltonian.

In contrast to the Foldy–Wouthuysen reduction, the direct Pauli reduction leads to a  $4 \times 4$  Hamiltonian  $H_I^P = T_0 H_I T_0^\dagger$  which is not block–diagonal in general. Thus the terms in  $S_{fi}$  involving negative–energy states will not vanish. Thus if we use  $H_I^{eff-P}$  consistently in a nonrelativistic scattering theory we will get an effective S–matrix  $S_{fi}^{eff-P}$  which is not equal to the (correct) relativistic result and which is therefore different from that obtained using  $H_I^{eff-FW}$ , by virtue of the  $\Phi_{0p}^{FW(-)}$  terms in eq. (46). In other words, using  $H_I^{eff-P}$  in a nonrelativistic approach amounts to omitting the negative–frequency contribution of the covariant calculation. One may expect that this approximation will lead to unacceptable results in situations where the original interaction Hamiltonian in the Dirac representation produces a significant coupling between upper and lower components.

We conclude from this that in cases in which both interactions are known relativistically, so that one can do a Foldy–Wouthuysen reduction on the complete interaction, as for example Compton scattering, use of  $H_I^{eff-FW}$  gives the correct result to given order in  $1/M$ , the same result as would be obtained in the complete relativistic theory. In contrast the Pauli reduction effective Hamiltonian  $H_I^{eff-P}$  gives different, and incorrect, results.

In many practical cases however one knows only one of the interactions relativistically. Thus for example in proton–proton bremsstrahlung or in

the electrodisintegration of nuclei the electromagnetic interaction is known relativistically but the strong interaction is not. In such cases the reduction is done on the known interaction alone, and one cannot generate the contact terms which involve both interactions. Thus it appears that for these cases one cannot prove rigorously that one effective Hamiltonian is more correct than another<sup>5</sup>.

To see how this works in a still somewhat simplified case, suppose that the interaction part of the Dirac Hamiltonian is given as before by harmonic interactions of the form  $H_I = H_{1a} + H_{1b}$ . Then by carrying out exactly the same steps used to evaluate the  $\Delta H \times \Delta H$  term in eq. (40) we find for the part of the full relativistic S-matrix which is first-order in each of the interactions

$$S_{fi} = -2\pi i \delta(E_f + \omega_b - E_i - \omega_a) \int d^3p \times \left[ \sum_{\text{spins}} \left\{ \frac{\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(+)} \rangle \langle \Phi_{0p}^{(+)} | H_{1a} | \Phi_{0i} \rangle}{E_f + \omega_b - E_p + i\epsilon} + \frac{\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(-)} \rangle \langle \Phi_{0p}^{(-)} | H_{1a} | \Phi_{0i} \rangle}{E_f + \omega_b + E_p - i\epsilon} \right\} + \text{c. t.} \right]. \quad (47)$$

We suppose that we do not know the interaction  $H_{1b}$  in a relativistic form. This means that we must approximate the full matrix elements  $\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(+)} \rangle$  and  $\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(-)} \rangle$  in some way. Typically one would neglect the latter and evaluate the former by using a nonrelativistic matrix element calculated, say, from a potential. Then one would reduce the known interaction  $H_{1a}$  to an effective Hamiltonian.

Using eqs. (35) and (36) one has  $H_{1a} = T_0^\dagger H_{1a}^P T_0$  in the Pauli reduction case and  $H_{1a} = T_0^\dagger H_{1a}^{FW} T_0 - \Delta H'_{1a}$  with  $\Delta H'_{1a} = i[S_{1a}, H_0] - \dot{S}_{1a}$  in the Foldy-Wouthuysen case. Substituting these into eq. (47) above and using eq. (41) to evaluate the  $\Delta H$  terms we find for the part in brackets

$$\left\{ \frac{\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(+)} \rangle \langle \Phi_{0p}^{FW(+)} | H_{1a}^{eff-FW} | \Phi_{0i}^{FW} \rangle}{E_f + \omega_b - E_p + i\epsilon} - \frac{\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(+)} \rangle \langle \Phi_{0p}^{FW(+)} | T_0 i S_{1a} T_0^\dagger | \Phi_{0i}^{FW} \rangle}{E_f + \omega_b + E_p - i\epsilon} \right\}$$

---

<sup>5</sup>Of course there may be other guiding principles, such as gauge invariance, which put a constraint on the form of the effective Hamiltonian.

$$- \langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(-)} \rangle \langle \Phi_{0p}^{FW(-)} | T_0 i S_{1a} T_0^\dagger | \Phi_{0i}^{FW} \rangle \Big\}. \quad (48)$$

The result using  $H_{1a}^{eff-P}$  is of the same form, except that the second term does not appear and  $H_{1a}^{eff-FW}$  is replaced by  $H_{1a}^{eff-P}$ .

The first term in this expression is what one would usually use in second-order perturbation theory in cases where only one of the interactions is known. It would lead to an approximate S-matrix,  $S_{fi}^{approx}$ , which differs from the full relativistic S-matrix by contact terms. Using an expansion analogous to eq. (5) and explicit forms for the leading parts of  $S_{1a}$  and  $T_0$  it is straightforward to show that the leading term in  $T_0 i S_{1a} T_0^\dagger$  is odd and of order  $\mathcal{O}_I/M$  whereas the first even term is down by one power of  $1/M$ . Thus the contact term involving the negative-energy intermediate states is likely to be the most important, whereas the one involving positive-energy intermediate states (which appears only for  $H_{1a}^{eff-FW}$ ) is down by  $1/M$ . One can see this explicitly in Compton scattering [23].

Above we showed that when both interactions are known  $H_I^{eff-P}$  fails to give the correct effective S-matrix because of the neglect of the terms involving negative-energy intermediate states. On the other hand  $H_I^{eff-FW}$  gives the correct answer, since it is constructed as the upper left-hand block of a Hamiltonian which does not connect positive- and negative-energy states, and since it contains the contact terms explicitly. Now when only one of the interactions is known the natural way of calculating  $S_{fi}^{approx}$  in second-order perturbation theory forces one to neglect these contributions for *both* effective Hamiltonians, and leads to an error which to leading order is the same in the two cases.

In most practical calculations, particularly when the unknown interaction is the strong interaction, second-order perturbation theory is not sufficient. In such cases, an example being proton-proton bremsstrahlung, one must sum the strong interaction to all orders. Thus one normally would replace  $\langle \Phi_{0f} | H_{1b} | \Phi_{0p}^{(+)} \rangle$  in eq. (48) by a full T-matrix. The cross term gives the interaction in the other order and there is an additional term, called double scattering in proton-proton bremsstrahlung, which has a T-matrix element on both sides of the matrix element of  $H_{1a}^{eff}$ . An alternative, but essentially equivalent approach which is often used is simply to take matrix elements of  $H_{1a}^{eff}$  between eigenstates of the free-plus-strong Hamiltonian.

Neither of these approaches will in general give the correct (relativistic)



answer, as long as the effective operator  $H_{1a}^{eff}$  is obtained as a reduction of the relativistic Hamiltonian between *free* states, because of inconsistency of the operator and the wave functions. In other words, approximations will be involved in getting from the full relativistic S-matrix to the effective nonrelativistic one. Regardless of the choice of  $H_{1a}^{eff}$ , contact terms will be missed.

It appears then that here, in contrast to the situation when both interactions are known relativistically, we cannot argue rigorously that  $H_I^{eff-FW}$  is any more correct than  $H_I^{eff-P}$ . At best it may be preferred simply because it is no worse an approximation when the interaction is incompletely known and is in principle the correct choice when the interaction is completely known.

## 5 Transformation of the S-Matrix under a Unitary Transformation

In the previous section we have explicitly shown to second-order perturbation theory that the S-matrix remains invariant under a unitary transformation of the states<sup>6</sup>. Here we will show this in general without making use of a perturbative approach. We start with the Dirac equation

$$i \frac{\partial \Psi(x)}{\partial t} = (H_0 + H_I(x, \xi)) \Psi(x) = H(x, \xi) \Psi(x), \quad (49)$$

where  $H_0 = \vec{\alpha} \cdot \vec{p} + \beta M$ ,  $x = (x^0, \vec{x})$ . We control the switching on and off of the interaction in the remote past and distant future through  $H_I(x, \xi) = H_I(x) e^{-\xi|t|}$ ,  $\xi \geq 0$ . The defining equation for the full propagator in the presence of interaction reads [2, 3]

$$\left( i \frac{\partial}{\partial t} - H_0 - H_I(x, \xi) \right) S_{F, H_I}(x, y) \beta = \delta^4(x - y), \quad (50)$$

---

<sup>6</sup>In fact we can even allow for more general transformations (see also ref. [29]). It turns out that the following proof is still correct if we demand  $T(t)$  to be an invertible transformation with the property of eq. (57). Note however that  $T_0$  remains a time-independent *unitary* transformation. One then only has to replace  $T^\dagger$  by  $T^{-1}$  in the proof.

and a perturbative solution satisfying the Feynman–Stückelberg boundary condition is given by

$$\begin{aligned} S_{F,H_I}(x,y) &= S_F(x-y) + \int d^4x_1 S_F(x-x_1)\beta H_I(x_1,\xi)S_F(x_1-y) \\ &+ \int d^4x_1 d^4x_2 S_F(x-x_2)\beta H_I(x_2,\xi)S_F(x_2-x_1)\beta H_I(x_1,\xi)S_F(x_1-y) \\ &+ \dots, \end{aligned} \quad (51)$$

where  $S_F(x-y)$  is the free Feynman propagator [2, 3]. For  $y_1^0 < x^0 < y_2^0$  the formal solution for  $\Psi(x)$  may be constructed as

$$\Psi(x) = i \int d^3y_1 S_{F,H_I}(x,y_1)\beta\Psi^{(+)}(y_1) - i \int d^3y_2 S_{F,H_I}(x,y_2)\beta\Psi^{(-)}(y_2), \quad (52)$$

where  $\Psi^{(+)}(y_1)$  and  $\Psi^{(-)}(y_2)$  are superpositions of positive- and negative-energy solutions of the free Dirac equation, respectively. If we specify the boundary conditions as

$$\begin{aligned} \lim_{x^0 \rightarrow -\infty} \Psi^{(+)}(x) &= \Psi_{0i}^{(+)}(x), \\ \lim_{x^0 \rightarrow \infty} \Psi^{(-)}(x) &= 0, \end{aligned} \quad (53)$$

where  $\Psi_{0i}^{(+)}(x)$  is a positive-energy eigenfunction of the free Dirac equation with eigenvalue  $E_i > 0$ , we can either describe the scattering of a particle or pair annihilation in the potential,

$$\begin{aligned} S_{fi} &= \lim_{\xi \rightarrow 0_+} \lim_{x^0 \rightarrow \infty} \int d^3x \Psi_{0f}^{(+)\dagger}(x)\Psi(x) \quad \text{for particle scattering,} \\ S_{fi} &= \lim_{\xi \rightarrow 0_+} \lim_{x^0 \rightarrow -\infty} \int d^3x \Psi_{0f}^{(-)\dagger}(x)\Psi(x) \quad \text{for pair annihilation,} \end{aligned} \quad (54)$$

with  $\Psi_{0f}^{(-)}$  a negative-energy solution of the free Dirac equation. Likewise, specifying the boundary conditions

$$\begin{aligned} \lim_{x^0 \rightarrow \infty} \Psi^{(-)}(x) &= \Psi_{0i}^{(-)}(x), \\ \lim_{x^0 \rightarrow -\infty} \Psi^{(+)}(x) &= 0, \end{aligned} \quad (55)$$

leads to a description of the scattering of an antiparticle or pair creation in the potential,

$$\begin{aligned} S_{fi} &= \lim_{\xi \rightarrow 0_+} \lim_{x^0 \rightarrow -\infty} \int d^3x \Psi_{0f}^{(-)\dagger}(x) \Psi(x) \quad \text{for antiparticle scattering,} \\ S_{fi} &= \lim_{\xi \rightarrow 0_+} \lim_{x^0 \rightarrow \infty} \int d^3x \Psi_{0f}^{(+)\dagger}(x) \Psi(x) \quad \text{for pair creation.} \end{aligned} \quad (56)$$

When defining the S-matrix elements of eq. (54) and (56), it is assumed that the system was in an eigenstate of  $H_0$  with positive energy in the remote past or will be in an eigenstate of  $H_0$  with negative energy in the distant future, respectively. It interacts with the potential at intermediate times and evolves under the influence of  $H_0$  in the distant future or remote past, respectively. The S-matrix is the mathematical idealization of extending the interaction over the complete t-axis [28]. However, it is important to realize that the limit  $\xi \rightarrow 0_+$  has to be taken at the end (see e. g. ref. [3], p 165 f).

If we introduce a unitary transformation  $T(x)$  which depends on the interaction Hamiltonian  $H_I(x, \xi)$ , or parts of it, with the property

$$\lim_{H_I(x, \xi) \rightarrow 0} T(x) = T_0, \quad (57)$$

where  $T_0$  is a time-independent unitary transformation, we will show that

$$S'_{fi} = S_{fi}, \quad (58)$$

provided the boundary conditions and  $S'_{fi}$  are defined in an analogous way as in eq. (53) - (56), e. g. for particle scattering

$$\begin{aligned} \lim_{x^0 \rightarrow -\infty} \Psi'^{(+)}(x) &= \Psi_{0i}'^{(+)}(x), \\ \lim_{x^0 \rightarrow \infty} \Psi'^{-}(x) &= 0, \\ S'_{fi} &= \lim_{\xi \rightarrow 0_+} \lim_{x^0 \rightarrow \infty} \int d^3x \Psi_{0f}'^{(+)\dagger}(x) \Psi'(x). \end{aligned} \quad (59)$$

In eq. (59)  $\Psi_{0i}'^{(+)}(x)$  and  $\Psi_{0f}'^{(+)}(x)$  are positive-energy eigenstates of  $H'_0 = T_0 H_0 T_0^\dagger$  and  $\Psi'(x)$  satisfies the Dirac equation

$$i \frac{\partial \Psi'(x)}{\partial t} = (H'_0 + H'_I(x, \xi)) \Psi'(x) = H'(x, \xi) \Psi'(x), \quad (60)$$

with  $H'(x, \xi) = T(x)H(x, \xi)T^\dagger(x) - iT(x)\partial T^\dagger(x)/\partial t$ . The extension to the other cases is straightforward. We then find for the S-matrix element  $S'_{fi}(\xi)$

$$S'_{fi}(\xi) = \lim_{x^0 \rightarrow \infty} \int d^3x \Psi_{0f}^{(+)\dagger}(x) T_0^\dagger T(x) \Psi(x). \quad (61)$$

For any arbitrarily small but finite  $\xi$  eq. (57) leads to

$$\lim_{x^0 \rightarrow \infty} T_0^\dagger T(x) = 1 \quad (62)$$

and we therefore obtain

$$S'_{fi}(\xi) = S_{fi}(\xi). \quad (63)$$

Taking the limit  $\xi \rightarrow 0_+$  in eq. (63) yields the desired result, eq. (58).

The above derivation is extremely simple, but it does not explicitly reveal the importance of  $H'(x, \xi)$  of eq. (60), in particular, it does not show why it is  $H'(x, \xi)$  and *not*  $T(x)H(x, \xi)T^\dagger(x)$  which enters the calculation of the primed S-matrix element. For that reason we provide a slightly more complicated derivation which, however, gives more insight into the role played by  $H'(x, \xi)$ . We once again only discuss the case of particle scattering. Using eq. (52) and the boundary condition of eq. (59) we may write  $S'_{fi}$  as

$$\begin{aligned} S'_{fi} &= i \lim_{\xi \rightarrow 0_+} \lim_{\substack{x^0 \rightarrow \infty \\ y^0 \rightarrow -\infty}} \int d^3x d^3y \Psi_{0f}^{(+)\dagger}(x) S'_{F, H'_I}(x, y) \beta \Psi_{0i}^{(+)}(x) \\ &= i \lim_{\xi \rightarrow 0_+} \lim_{\substack{x^0 \rightarrow \infty \\ y^0 \rightarrow -\infty}} \int d^3x d^3y \Psi_{0f}^{(+)\dagger}(x) T_0^\dagger S'_{F, H'_I}(x, y) \beta T_0 \Psi_{0i}^{(+)}(x), \end{aligned} \quad (64)$$

where  $S'_{F, H'_I}(x, y)$  is defined by an equation analogous to eq. (50), namely

$$\left( i \frac{\partial}{\partial t} - H'_0 - H'_I(x, \xi) \right) S'_{F, H'_I}(x, y) \beta = \delta^4(x - y). \quad (65)$$

Clearly, at this point one realizes that  $H'(x)$  rather than  $T(x)H(x, \xi)T^\dagger(x)$  is the relevant operator in the defining equation for the full propagator. It may easily be shown that

$$S'_{F, H'_I}(x, y) \beta = T(x) S_{F, H_I}(x, y) \beta T^\dagger(y) \quad (66)$$

solves the equation of motion for  $S'_{F,H_I}$ , eq. (65), provided  $S_{F,H_I}$  satisfies eq. (50). We then insert eq. (66) into eq. (64) to obtain

$$S'_{fi} = i \lim_{\xi \rightarrow 0+} \lim_{\substack{x^0 \rightarrow \infty \\ y^0 \rightarrow -\infty}} \int d^3x d^3y \Psi_{0f}^{(+)\dagger}(x) T_0^\dagger T(x) S_{F,H_I}(x, y) \beta T^\dagger(y) T_0 \Psi_{0i}^{(+)}(x). \quad (67)$$

Once again we make use of eq. (57) for any arbitrarily small but finite  $\xi$ ,

$$\lim_{y^0 \rightarrow -\infty} T^\dagger(y) T_0 = 1 = \lim_{x^0 \rightarrow \infty} T_0^\dagger T(x), \quad (68)$$

and obtain

$$\begin{aligned} S'_{fi}(\xi) &= i \lim_{\substack{x^0 \rightarrow \infty \\ y^0 \rightarrow -\infty}} \int d^3x d^3y \Psi_{0f}^{(+)\dagger}(x) S_{F,H_I}(x, y) \beta \Psi_{0i}^{(+)}(x) \\ &= S_{fi}(\xi), \end{aligned} \quad (69)$$

which is identical with eq. (63).

In conclusion, we have provided two nonperturbative arguments for the S-matrices to be equal provided the states are related by a unitary transformation.

## 6 Summary and Conclusions

We have addressed in this paper the question of how one obtains effective nonrelativistic Hamiltonians from known, or partially known, relativistic interactions. Such Hamiltonians are important because they allow the use of nonrelativistic formalisms, and make connection with interactions which may be known only in some nonrelativistic approximation, and yet at the same time incorporate relativistic corrections to some order in  $1/M$ .

We discussed two different ways of obtaining such effective  $2 \times 2$  Hamiltonians. The first is the Foldy–Wouthuysen interaction which is obtained by a unitary time-dependent transformation of the states. This approach decouples the upper from the lower components order by order in  $1/M$  and so, by projecting out the upper left-hand block, leads naturally to an effective Hamiltonian  $H_I^{eff-FW}$  to be used with two-component positive-energy wave functions.

The second method consists of making a two-component Pauli reduction of the matrix element of the interaction Hamiltonian in the Dirac representation between free positive-energy spinors. The result is then regarded as an effective Hamiltonian  $H_I^{eff-P}$  to be used between nonrelativistic Pauli wave functions. This approach is equivalent to making a unitary transformation of the states using the time-independent transformation  $T_0$  and then projecting out the upper left-hand block to get  $H_I^{eff-P}$ . Usually the result is expanded in  $1/M$ , typically to order  $1/M^2$  or  $1/M^3$ , though in some applications this step is omitted, i. e. the complete expression without a  $1/M$  expansion is used. Both of these methods for obtaining effective Hamiltonians and introducing relativistic corrections have been used for a variety of processes in the literature.

It has been tacitly assumed in the literature that both approaches yield the same results, i. e. that it makes no difference which of the two effective Hamiltonians is used. We have seen that this is *in general* not true and that in fact it does make a difference which is used, at least in the case when the initial relativistic interaction is known completely.

For first-order matrix elements differences in the results obtained using the two different effective Hamiltonians arise only when at least one state is off-energy-shell, or in other words only when energy is not conserved, as the difference of the matrix elements calculated with different effective Hamiltonians is proportional to the sum of initial energies minus the sum of the final energies at a vertex.

However, in higher-than-first-order, time-ordered perturbation theory, energy is not conserved at each individual vertex. Furthermore intermediate negative-energy states contribute and so we might expect effects from the projection eliminating the effects of the negative-energy states which is used in getting the effective Hamiltonians from the relativistic ones. We found in second-order perturbation theory that the effective S-matrices, obtained by using the  $2 \times 2$  effective Hamiltonians in usual nonrelativistic scattering theory were in fact different. The result obtained using  $H_I^{eff-FW}$  agreed with the full relativistic S-matrix while that obtained with  $H_I^{eff-P}$  did not.

At first glance this seems inconsistent with the perception that the observables should not change under a unitary transformation. As we showed, both in second-order perturbation theory and via a general argument, the full relativistic S-matrix is invariant under such time-dependent unitary transformations, of which  $T(t)$  leading to  $H_I^{FW}$  and  $T_0$  leading to  $H_I^P$  are examples.

This apparent conflict was resolved via an understanding of the additional approximations necessary in going from the full relativistic S-matrix to the effective S-matrix calculated nonrelativistically. To get  $S_{fi}^{eff}$  one must neglect the negative-energy intermediate states. Since  $H_I^{FW}$  is block-diagonal and does not connect positive- to negative-energy states anyway, this approximation has no effect and in this case the effective S-matrix  $S_{fi}^{eff-FW}$  is the same as the full relativistic one. In contrast  $H_I^P$  does connect negative- and positive-energy states. Thus the additional approximation needed to get  $H_I^{eff-P}$ , namely the projection for positive energies which picks out just the upper left-hand block, leads to the neglect of some nonzero terms and thus to an effective S-matrix  $S_{fi}^{eff-P}$  which is different from the full relativistic S-matrix.

The results just described apply only to cases where the full relativistic interaction is known, as for example Compton scattering or pion photoproduction. In many practical situations only part of the relativistic Hamiltonian is known. Typically, such a situation may involve a nonrelativistic potential approach for the strong interaction and at the same time a relativistic treatment of the electromagnetic interaction (including the anomalous magnetic moment and possibly on-shell form factors).

In such cases the situation is ambiguous. Lack of knowledge of one of the interactions makes it impossible to calculate the contact terms explicitly and hence they must be neglected whichever effective Hamiltonian is used. Use of  $H_I^{eff-FW}$  requires one to drop an additional contact term, which is however of order  $1/M$  times the main contact terms which are dropped in all cases. For some processes gauge invariance allows one to construct non-unique contact terms which can be added in by hand. However in general there may be numerical differences when different effective Hamiltonians are used. Unfortunately it does not seem to be possible to determine in a rigorous way which is best for these situations, though there may be a philosophical predilection to use the Foldy-Wouthuysen approach even in the absence of full information as that is the correct approach when the interaction is fully known.

In conclusion then, our analysis indicates that when the interaction is fully known relativistically use of the Foldy-Wouthuysen reduction in order to incorporate relativistic corrections to a nonrelativistic treatment leads to correct results. In contrast using the matrix element of the interaction Hamiltonian in the Dirac representation between free positive-energy states

as an effective nonrelativistic Hamiltonian gives incorrect results. However when only part of the interaction is known, as is the case in many practical situations, both methods miss the leading order contact terms. Thus in this situation one cannot argue rigorously that one should be preferred over the other.

## Acknowledgements

This work was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank J. H. Koch for useful comments.

## References

- [1] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78** (1950) 29
- [2] J. D. Bjorken and S. D. Drell, Relativistic Quantum Mechanics, McGraw-Hill, New York, 1964
- [3] C. Itzykson and J.-B. Zuber, Quantum Field Theory, McGraw-Hill, New York, 1980
- [4] M. V. Barnhill III, Nucl. Phys. **A131** (1969) 106
- [5] J. L. Friar, Phys. Lett. **B59** (1975) 145
- [6] H. Hyuga and M. Gari, Nucl. Phys. **A274** (1976) 333
- [7] K. Ohta and M. Ichimura, Nucl. Phys. **A491** (1989) 509
- [8] H. Göller and H. Arenhövel, Johannes Gutenberg-Universität, Mainz, preprint MKPH-T-92-3 (1992)
- [9] J. Dubach, J. H. Koch and T. W. Donnelly, Nucl. Phys. **A271** (1976) 279
- [10] W. Fabian and H. Arenhövel, Nucl. Phys. **A314** (1979) 253
- [11] J.-F. Mathiot, Phys. Rep. **173** (1989) 63



- [12] W. M. Alberico, T. W. Donnelly and A. Molinari, Nucl. Phys. **A512** (1990) 541
- [13] E. Hadjimichael, G. I. Poulis and T. W. Donnelly, Phys. Rev. **C45** (1992) 2666
- [14] J. L. Friar, Ann. Phys. **104** (1977) 380
- [15] K. Ohta, J. Phys. G **14** (1988) 449
- [16] R. L. Workman and Harold W. Fearing, Phys. Rev. **C34** (1986) 780; Harold W. Fearing, Nucl. Phys. **A463** (1987) 95
- [17] V. Herrmann and K. Nakayama, Phys. Rev. C **46** (1992) 2199 and references cited therein; K. Nakayama, private communication
- [18] M. Jetter and H. V. von Geramb, Proceedings of the International Conference on Nuclear Physics, Wiesbaden, July, 1992
- [19] V. R. Brown, P. L. Anthony and J. Franklin, Phys. Rev. C **44** (1991) 1296
- [20] A. Katsogiannis and K. Amos, Phys. Rev. C **47** (in press)
- [21] M. M. Nieto, Phys. Rev. Lett. **38** (1977) 1042
- [22] T. Goldman, Phys. Rev. **D15** (1977) 1063
- [23] S. Scherer, G. I. Poulis and H. W. Fearing, TRIUMF preprint 93-13 (1993)
- [24] K. W. McVoy and L. Van Hove, Phys. Rev. **125** (1962) 1034
- [25] W. Pauli, Rev. Mod. Phys. **13** (1941) 203
- [26] S. Weinberg, Phys. Rev. **150** (1966) 1313
- [27] F. Halzen and A. D. Martin, Quarks and Leptons, John Wiley & Sons, New York, 1984
- [28] N. N. Bogoliubov and D. V. Shirkov, Quantum Fields, Benjamin/Cummings, Reading, 1983

- [29] S. Kamefuchi, L. O’Raifeartaigh and A. Salam, Nucl. Phys **28** (1961) 529

**Figure 1:** Cross section and analyzing power for proton–proton bremsstrahlung for an incident laboratory energy of 280 MeV and a coplanar equal angle geometry for the outgoing protons, plotted versus the photon angle. The calculations were done using the full Bonn potential and the formalism of ref. [16] except that Coulomb corrections are not included. The solid (dotted) curve corresponds to using  $H_I^{eff-FW}$  ( $H_I^{eff-P}$ ) for the electromagnetic interaction. Note the suppressed zeros on both axes.